

Classification of very high resolution SAR images of urban areas by dictionary-based mixture models, copulas and Markov random fields using textural features

Aur lie Voisin^{a#}, Gabriele Moser^b, Vladimir A. Krylov^{a,c}, Sebastiano B. Serpico^b
and Josiane Zerubia^a

^a EPI Ariana, CR INRIA Sophia Antipolis M diterran e, 2004, Route des Lucioles, B.P.93,
F-06902 Sophia Antipolis (France);

^b Dept. of Biophysical and Electronic Engineering (DIBE), University of Genoa, Via Opera
Pia 11a, I-16145, Genoa (Italy);

^c Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State
University, 119991 Leninskie Gory, Moscow (Russia).

ABSTRACT

This paper addresses the problem of the classification of very high resolution (VHR) SAR amplitude images of urban areas. The proposed supervised method combines a finite mixture technique to estimate class-conditional probability density functions, Bayesian classification, and Markov random fields (MRFs). Textural features, such as those extracted by the greylevel co-occurrence method, are also integrated in the technique, as they allow to improve the discrimination of urban areas. Copulas are applied to estimate bivariate joint class-conditional statistics, merging the marginal distributions of both textural and SAR amplitude features. The resulting joint distribution estimates are plugged into a hidden MRF model, endowed with a modified Metropolis dynamics scheme for energy minimization. Experimental results with COSMO-SkyMed and TerraSAR-X images point out the accuracy of the proposed method, also as compared with previous contextual classifiers.

Keywords: Synthetic Aperture Radar (SAR), Supervised classification, Urban areas, Textural features, Copulas, Markov Random Fields

1. INTRODUCTION

Synthetic aperture radar (SAR)¹ is known to be unaffected by sun-illumination, and almost not influenced by atmospheric conditions. Recent improvements of spaceborne SAR currently enable acquisitions of Very High Resolution (VHR) data (up to metric resolution) with a very short revisit time (up to 12 hours). The acquisitions may be either single- or multi-polarized, thus highlighting different aspects of a same ground area. In this respect, SAR imagery offers a huge potential for risk management² e.g. by allowing land-use and/or land-cover mapping or detection of areas damaged by a disaster event. In the framework of the assessment of environmental risk, we address here the problem of classifying SAR images of urban areas, a specifically interesting typology given the fact that it is strategic and critical for population risks.

A typical difficulty of SAR image processing is posed by speckle¹. A further issue is the spatial heterogeneity of urban areas. For this reason, classification methods developed for VHR optical images including urban areas³ are often not adaptable for SAR. Lots of methods have already been considered for SAR amplitude image classification in both supervised and unsupervised contexts, based on maximum likelihood⁴, triplet Markov fields⁵, neural networks⁶, bags-of-features⁷ and other methods⁸. The classification considered in this paper is developed in a supervised context. It is an extension of the method proposed by Moser and al.⁹ for supervised classification of VHR single- and dual-polarized SAR areas and tested for wet soil mapping in application to epidemiology. The method models the statistics of SAR amplitude probability density function (PDF) using a

[#] E-mail: aurelie.voisin@sophia.inria.fr

dictionary-based stochastic expectation maximization (DSEM) approach, i.e. the statistics of SAR amplitudes are considered as mixtures of K parametric components that are automatically drawn from a dictionary of SAR-specific PDFs.^{10,11} When separately applied to each class considered in the VHR SAR image, DSEM represents a natural model for heterogeneous scenarios, leading to a mixture estimate where distinct components may be interpreted as the contributions of different ground materials present in each relevant class. This makes the algorithm robust with respect to possibly complicated shapes of class histograms. The marginal DSEM PDF estimates are used to model the joint PDF of the SAR amplitude image and of an extracted textural feature by the use of copulas¹², which enables to take into account both their contributions. To proceed to contextual image classification, we combine the resulting joint PDF estimate with a Markov random field (MRF) approach¹³. MRFs represent a general family of probabilistic image models that provide a convenient and consistent way to characterize context-dependent data¹⁴.

This paper is organized as followed. In Section 2, we present an overview of the proposed classification algorithm. In Section 3, we focus on DSEM. In Section 4, we introduce the textural features used to discriminate urban areas, and how they are plugged into the DSEM-MRF algorithm by copulas. Section 5 reports experimental results of the application of the developed method to COSMO-SkyMed and TerraSAR-X images. We conclude the paper in Section 6.

2. CLASSIFICATION METHOD OVERVIEW

The DSEM-MRF method consists of three steps: SAR amplitude PDF estimation, MRF modeling and classification.

2.1 Probability density function estimation

This step aims to model the PDF distribution for each class m considered for classification, $m \in [1; M]$, given a training map, by using a dictionary-based stochastic expectation maximization. For each class, the mixture PDFs $p_m(z|\omega_m)$ are estimated following a finite mixture model¹⁵ for the independent distribution of greylevels

$$p_m(z|\omega_m) = \sum_{i=1}^K P_{mi} p_{mi}(z|\theta_{mi}), \quad (1)$$

where z is a greylevel, $z \in [0; Z - 1]$, and ω_m is the m^{th} class. P_{mi} are mixing proportions, such that for a given m , $\sum_{i=1}^K P_{mi} = 1$ with $0 \leq P_{mi} \leq 1$. The various mixture PDF models $p_{mi}(z|\theta_{mi})$ are selected in a predefined dictionary (Tab. 1). This PDF selection, as well as the calculation of the mixing proportions, and the estimation of the parameter K are further detailed in Sec. 3.

The use of DSEM instead of single parametric PDF models proves to be advantageous when modeling heterogeneous classes, e.g. urban areas are composed of different materials (roofs, asphalt, cars) clearly visible in very high resolution. Moreover, whether in homogeneous or heterogeneous areas, DSEM can be viewed as an efficient tool for choosing the best single PDF model from a predefined PDF dictionary.

As described in Sec.4, DSEM is used to model the marginal PDF estimates of both textural features and SAR image amplitudes, then combined by copulas so as to lead to joint bivariate PDFs.

2.2 Markov Random Field

To gain robustness against speckle, and take into account the influence of neighboring pixels (contextual information), we consider a Markov model. By the Gibbs property and the Hammersley-Clifford theorem¹⁶, we can define a local characteristic (conditional probability) for each class $m \in [1; M]$

$$p(x_s = \omega_m | x^{(s)}) = \frac{\exp(-H(x_s = \omega_m | x^{(s)}))}{\sum_{j=1}^M \exp(-H(x_s = \omega_j | x^{(s)}))} \quad (2)$$

where s is the current site ($s \in S$), x_s the corresponding class label, $x^{(s)}$ the configuration outside the site s such that $x^{(s)} = \{x_t, t \neq s, t \sim s\}$ and $t \sim s$ means that t and s are neighboring pixels. We deal here with an

anisotropic second order neighborhood; only cliques C of size 2 are considered. The MRF energy function H , given the conditional PDFs (1) and no prior information about the proportions of classes on the testing image, is expressed with only one parameter $\beta > 0$:

$$H(\omega_m|z, \beta) = \sum_{s' \in S} \left[-\log p_m(z|\omega_m) - \beta \sum_{s: \{s, s'\} \in C} \delta_{x_s=x_{s'}} \right] \quad \text{with } \delta_{x_s=x_{s'}} = \begin{cases} 1, & \text{if } x_s = x_{s'} \\ 0, & \text{otherwise} \end{cases}. \quad (3)$$

Further information about MRFs and hidden MRFs can be found in Ref.17.

According to Eqs. (1)-(3), the considered MRF model includes only one parameter β to estimate. This is due to the fact that all the parameter estimation for (1) is already incorporated into DSEM. With the knowledge of a training map, we can easily estimate β , by maximizing the pseudo-likelihood PL , defined with the local characteristics of Eq. (2):

$$\log PL(x|\beta) = \log \left[\prod_{s \in S} p(x_s = \omega_m | x^{(s)}, \beta) \right]. \quad (4)$$

To optimize this function, a simulated annealing algorithm turns out to be effective¹⁸. If necessary, this procedure can be followed by the method suggested in Ref. 19, combining Metropolis-Hasting algorithm with gradient method.

Preliminary experiments showed that for a correct β estimation, the training ground truth must be exhaustive, or at least sufficiently representative of class-transition areas. This is consistent with the role of this parameter in tuning the probability of spatial class transitions.

2.3 Classification

In order to generate the output classification map, the energy function H (Eq. (3)) is minimized by the modified Metropolis dynamics (MMD) algorithm that is usually an effective tradeoff between accuracy and computational burden⁹. It is a compromise between the deterministic Iterated Conditional Modes algorithm¹⁷, which is fast and finds a local minimum, and Simulated Annealing¹⁸, that is much slower, but may find the global minimum.

3. MARGINAL PDF ESTIMATION BY DSEM

In this part, we describe more precisely the algorithm used for estimating the PDF, considering a finite mixture model (Eq. (1)). This approach was proposed in Ref. 10, then developed and validated on VHR SAR in Ref. 11.

Each mixture component is automatically chosen in a predefined dictionary $D = \{f_1, f_2, f_3, f_4\}$ (see Table 1) of four SAR-specific PDFs¹¹.

DSEM integrates the estimation of the number of parameters K , the selection of the optimal model for each component p_{mi} , the calculation of their corresponding parameters θ_{mi} and the mixture proportions P_{mi} . We deal here with incomplete data because we do not know to which of the K components the i^{th} greylevel belongs. Thus, we introduce missing labels s_i such that complete data are defined by the set $\{(z_i, s_i), i \in [1; N]\}$. An efficient algorithm dealing with this data incompleteness is the stochastic expectation maximization²⁰ (SEM).

Instead of adopting Maximum Likelihood (ML) estimates as in the classical SEM scheme, which is not always feasible for the considered SAR PDFs¹¹ (e.g., for Nakagami distribution), we adopt here the Method of Log-Cumulants^{10,21} (MoLC). By analogy to the Laplace transform in moment generating function²², MoLC uses the Mellin transform²³ and allows stating a set of equations relating the unknown parameters of a given parametric model with one or several log-cumulants:

$$\begin{cases} \kappa_1 = E\{\ln u\} \\ \kappa_2 = \text{Var}\{\ln u\} \\ \kappa_3 = E\{(\ln u - \kappa_1)^3\} \end{cases}. \quad (5)$$

These equations have one solution for any observed log-Cumulants for all PDFs in D (see Tab. 1).

For each class m , each iteration t of DSEM goes as follows:

Table 1. PDFs and MoLC equations for the parametric families included in the considered dictionary D . Here, $\Gamma(\cdot)$ is the Gamma function²³, $\Psi(\cdot)$ the Digamma function²³ and $\Psi(\nu, \cdot)$ the ν^{th} order polygamma function²³.

Family	Probability density function	MoLC equations
Log-normal	$f_1(r) = \frac{1}{\sigma r \sqrt{2\pi}} \exp \left[-\frac{(\ln r - m)^2}{2\sigma^2} \right]$	$\kappa_1 = m$ $\kappa_2 = \sigma^2$
Weibull	$f_2(r) = \frac{\eta}{\mu^\eta} r^{\eta-1} \exp \left[-\left(\frac{r}{\mu}\right)^\eta \right]$	$\kappa_1 = \ln \mu + \Psi(1)\eta^{-1}$ $\kappa_2 = \Psi(1, 1)\eta^{-2}$
Nakagami	$f_3(r) = \frac{2}{\Gamma(L)} (\lambda L)^L r^{2L-1} \exp(-\lambda L r^2)$	$2\kappa_1 = \Psi(L) - \ln \lambda L$ $4\kappa_2 = \Psi(1, L)$
Generalized Gamma	$f_4(r) = \frac{\nu}{\sigma \Gamma(\kappa)} \left(\frac{r}{\sigma}\right)^{\kappa\nu-1} \exp \left\{ -\left(\frac{r}{\sigma}\right)^\nu \right\}$	$\kappa_1 = \Psi(\kappa)/\nu + \ln \sigma$ $\kappa_2 = \Psi(1, \kappa)/\nu^2$ $\kappa_3 = \Psi(2, \kappa)/\nu^3$

- **E-step:** compute, for each greylevel z and i^{th} component, the posterior probability estimates corresponding to the current PDF estimates, i.e. $z \in [0; Z-1]$, $i \in [1; K^t]$:

$$\tau_i^t(z) = \frac{P_i^t p_i^t(z)}{\sum_{j=1}^{K^t} P_j^t p_j^t(z)},$$

where $p_i^t(\cdot)$ is the current estimate for the i^{th} component;

- **S-step:** sample the label $s^t(z)$ of each greylevel z according to the current estimated posterior probability distribution $\{\tau_i^t(z) : i \in [1; K^t]\}$, $z \in [0; Z-1]$;
- **MoLC-step:** for the i^{th} mixture component, compute the following histogram-based estimates of the mixture proportion and the first three log-cumulants:

$$P_i^{t+1} = \frac{\sum_{z \in Q_{it}} h(z)}{\sum_{z=0}^{Z-1} h(z)}, \quad \kappa_{1i}^t = \frac{\sum_{z \in Q_{it}} h(z) \ln z}{\sum_{z \in Q_{it}} h(z)}, \quad \kappa_{bi}^t = \frac{\sum_{z \in Q_{it}} h(z) (\ln z - \kappa_{1i}^t)^b}{\sum_{z \in Q_{it}} h(z)},$$

where $b = 2, 3$, $h(z)$ is the image histogram, Q_{it} is the set of greylevels assigned to the i^{th} component; then, solve the corresponding MoLC equations (see Table 1) for each parametric family $f_j(\cdot|\theta_j)$ ($\theta_j \in A_j$) in the dictionary, thus computing the resulting MoLC estimate θ_{ij}^t , $i = 1, \dots, K^t$, $j = 1, \dots, M$;

- **K-step:** for all i , $i = 1, \dots, K^t$: if P_i^{t+1} is below a given threshold, eliminate the i^{th} component and update K^{t+1} ;
- **Model Selection-step:** for the i^{th} mixture component, compute the log-likelihood of each estimated PDF $f_j(\cdot|\theta_{ij}^t)$ according to the data assigned to the i^{th} component:

$$L_{ij}^t = \sum_{z \in Q_{it}} h(z) \ln f_j(z|\theta_{ij}^t)$$

and define $p_i^{t+1}(\cdot)$ as the estimated PDF $f_j(\cdot|\theta_{ij}^t)$ yielding the highest value of L_{ij}^t , $i = 1, \dots, K^{t+1}$, $j = 1, \dots, 4$.

4. COMBINATION WITH TEXTURAL FEATURES

Textural features often turn out to be discriminant with respect to urban areas (see Fig. 1). Moreover, the use of copulas allows taking into account the contribution of two input features by the calculation of a bivariate joint PDF. The flexibility of DSEM, granted by its essentially nonparametric formulation, makes it feasible to estimate the marginal PDFs of both the amplitude and the textural features. Thanks to Sklar's theorem¹²,

copula functions allow a joint bivariate PDF to be modeled, given the related marginal PDFs. This joint PDF is directly plugged into the hidden MRF model.

Two kinds of textural features were generated, using the greylevel co-occurrence matrix²⁴ (GLCM) method and the semivariogram approach²⁵. They are extracted from the original SAR image, leading to textural representations (see Fig. 1).

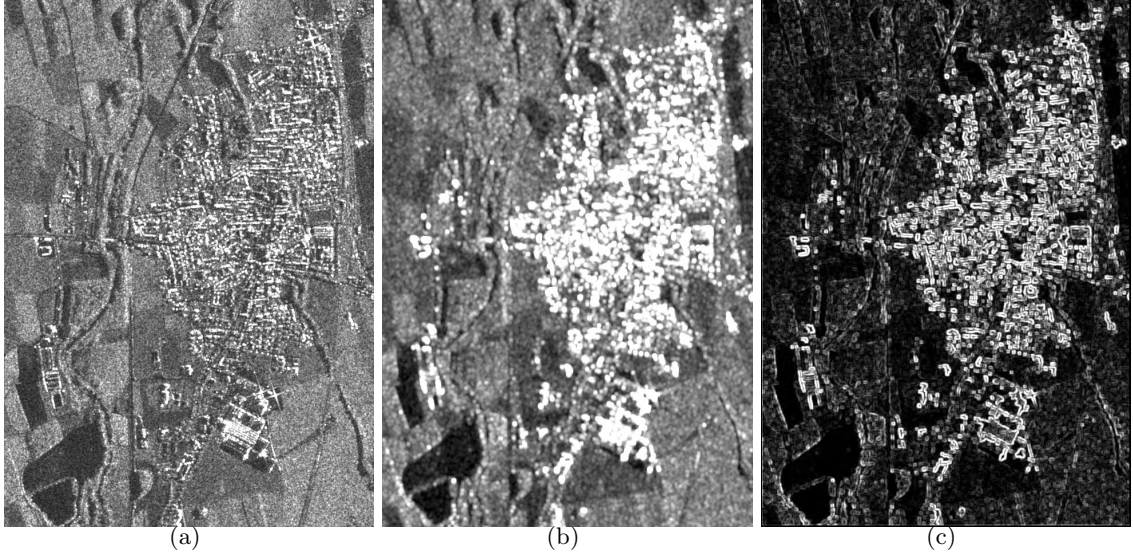


Figure 1. (a): Original SAR image (**CSK1**), Cavallermaggiore (Italy) (COSMO-SkyMed,©ASI); (b): Textural feature extracted with semivariogram; (c): Textural feature extracted with GLCM. The urban areas are well discriminated by these features.

4.1 GLCM and Semivariogram

The greylevel co-occurrence matrix G is a square $Z \times Z$ matrix, Z being the number of greylevels in the image, which describes the joint statistics of the greylevels of different pixels as a function of their reciprocal locations. Typically, the element (i, j) of the matrix is the probability $g(i, j)$ that a pixel with value i is adjacent to a pixel with value j . The adjacency can be defined in each of the possible directions (horizontal, vertical, left and right diagonals). In our case, we considered horizontal adjacency, with an offset equals to 1, i.e the matrix is filled by considering a reference pixel, and the pixel located to its right. Among various textures that can be extracted from the GLCM matrix, we use the variance, which usually discriminates well urban areas.

The *semivariogram* describes the spatial properties of an image, in particular the degree of spatial dependence of a spatial random field assessed in terms of second-order statistics. If s and t are two adjacent pixels (separated by a distance-offset h), the semivariogram is defined as the expected square increment of the values between s and t :

$$\gamma(s, t) = \frac{E[|z(s) - z(t)|^2]}{2} \quad (6)$$

Empirically, the semivariogram is expressed as:

$$\hat{\gamma}(h) = \frac{\sum_{(i,j) \in N(h)} |z_i - z_j|^2}{N(h)} \quad (7)$$

where i, j are adjacent pixels separated by an offset h , $N(h)$ denotes the set of pairs of observations and z_i, z_j their corresponding greylevels.

Both approaches to texture extraction are applied on a moving-window basis. Typically, we choose windows of size $w = 5$. Preliminary experiments (not reported for brevity) confirmed that this choice provided more accurate results (at least, on the considered data sets) than other window sizes.

4.2 Copula theory

The aim of using copulas is to compensate the fact that, contrary to marginal PDFs modeling, few models were proposed for modeling joint probabilities of SAR images (e.g. in dual-pol case) and no model is available for the joint statistics of SAR amplitudes and textures. We only focus here on the 2D case, but copulas theory can be generalized to multivariate cases.

2D copula is a bivariate joint distribution defined on $[0, 1]^2$ such that marginal distributions are uniform on $[0, 1]$. Specifically, a *bivariate copula* is a function $C : [0, 1]^2 \rightarrow [0, 1]$, which satisfies the following properties:

1. both marginals are uniformly distributed on $[0, 1]$;
2. for every u, v in $[0, 1]$: $C(u, 0) = C(0, v) = 0$, and $C(u, 1) = u$, $C(1, v) = v$;
3. for every $u_1 \leq u_2$, $v_1 \leq v_2$ in $[0, 1]$: $C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0$.

The importance of copulas in statistics is explained by Sklar's theorem¹², which states the existence of a copula C , that models the joint distribution function H of arbitrary random variables Z_1 and Z_2 with Cumulative Distribution Functions (CDFs) F and G :

$$H(z_1, z_2) = C(F(z_1), G(z_2)), \quad (8)$$

for all z_1, z_2 in \mathbb{R} . If F and G are continuous, then C is unique.

Taking the derivative in (8) over the two continuous random variables z_1 and z_2 with PDFs f and g , we obtain the joint PDF distribution:

$$h(z_1, z_2) = f(z_1)g(z_2) \frac{\partial^2 C}{\partial z_1 \partial z_2}(F(z_1), G(z_2)). \quad (9)$$

To find the best fitting copula C , we consider a dictionary of 5 copulas: Clayton, Ali-Mikhail-Haq, Frank, Marchal-Olkin and Farlie-Gumbel-Morgenstern. This choice of copulas is able to model a considerable variety of dependence structures²⁶. Each considered copula involves one parameter θ . We summarize the information about the copulas used in this work in Tab. 2.

Table 2. Considered copulas: Clayton, Ali-Mikhail-Haq (AMH), Frank, Marchal-Olkin (Marchal) and Farlie-Gumbel-Morgenstern (FGM) defined by their $C(u, v)$, with $\theta(\tau)$, τ -intervals and θ -intervals.

Copula	$C(u, v)$	$\theta(\tau)$	τ interval	θ interval
Clayton	$(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$	$\theta = \frac{2\tau}{1-\tau}$	$\tau \in]0; 1]$	$\theta \in]0; +\infty]$
AMH	$\frac{uv}{1-\theta(1-u)(1-v)}$	$\tau = \frac{3\theta-2}{3\theta} - \frac{2}{3} \left(1 - \frac{1}{\theta}\right)^2 \ln(1-\theta)$	$\tau \in [-0.182; \frac{1}{3}]$	$\theta \in [-1; +1]$
Frank	$-\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right)$	$\tau = 1 - \frac{4}{\theta^2} \int_0^{\theta} \frac{t}{e^{-t} - 1} dt$	$\tau \in [-1; 1] \setminus \{0\}$	$\theta \neq 0$
Marchal	$\min(u^{1-\theta} v, uv^{1-\theta})$	$\theta = \frac{2\tau}{\tau+1}$	$\tau \in [0; 1]$	$\theta \in [-1; +1]$
FGM	$\exp(uv(1+\theta(1-u)(1-v)))$	$\theta = \frac{9\tau}{2}$	$\tau \in [-2/9; 2/9]$	$\theta \in [-1; +1]$

To estimate θ and the best fitting copula, we use the relation between θ and Kendall's τ , which is a ranking correlation coefficient¹². By definition, *Kendall's* τ is a concordance-discordance measure between two independent realizations (Z_1, Z_2) and (\hat{Z}_1, \hat{Z}_2) : $\tau = P\{(Z_1 - \hat{Z}_1)(Z_2 - \hat{Z}_2) > 0\} - P\{(Z_1 - \hat{Z}_1)(Z_2 - \hat{Z}_2) < 0\}$. Given two realizations $z_{1,l}$ and $z_{2,l}$ ($l \in [1; N]$), the empirical estimator of Kendall's τ is:

$$\hat{\tau} = \frac{\sum_{l=1}^{N-1} \sum_{k=l+1}^N z_{1,lk} z_{2,lk}}{\binom{N}{2}}, \quad \text{where } z_{n,lk} = \begin{cases} 1, & \text{if } z_{n,l} \leq z_{n,k} \\ -1, & \text{otherwise} \end{cases}, \quad \text{for } n = 1, 2. \quad (10)$$

The general connection between Kendall's τ and the copula C associated with $H(z_1, z_2)$ is obtained by integrating the definition of τ over the distribution of (\hat{Z}_1, \hat{Z}_2) :

$$\tau = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1. \quad (11)$$

Once the estimate $\hat{\tau}$ (10) is computed, we plug it in place of τ in (11), so as to get parameter estimates $\hat{\theta}$ (see Tab. 2).

In order to estimate the choice of the copula in the considered dictionary, first, we discard the copulas for which the current estimate $\hat{\tau}_m$ is outside the τ -relevance interval (Tab. 2). Then, for each m , we choose the best fitting copula according to the highest p -value in Pearson Chi-square test-of-fitness (PCS)²⁷. The null hypothesis in PCS is that the sample frequencies $C_c(F_{1m}(u_1), F_{2m}(u_2))$ ($m = 1, \dots, M$, $c = 1, \dots, 5$), where (u_1, u_2) are the observed data, and F_{1m} and F_{2m} are the distribution estimates corresponding to the marginal DSEM density estimates, are consistent with the theoretical frequencies for the copula c : $C_c(v_1, v_2)$.

More specifically, in our case, each cumulative distribution function is easily formulated thanks to the knowledge of the PDF parameters (Eq. (1)), computed in the DSEM step:

$$F_m(z|\omega_m) = \sum_{i=1}^K P_{mi} F_{mi}(z|\theta_{mi}). \quad (12)$$

$F_{mi}(z|\theta_{mi})$ is the CDF, defined as the integral over $]-\infty; z]$ of the corresponding PDF $p_{mi}(z|\theta_{mi})$. Hence, once the joint PDF is calculated with the knowledge of the SAR image PDF and the textural feature PDF, the classification can be performed as explained in Sec. 2.2.

5. EXPERIMENTAL RESULTS

The proposed method was experimentally validated with COSMO-SkyMed and TerraSAR-X data. The following SAR amplitude images were used here to illustrate the results:

- **TSX1:** HH Polarized TerraSAR-X, SpotLight (8.2 m ground resolution), ellipsoid corrected, 4-look image acquired over Germany, city of Rosenheim; (©Infoterra, 2008); see Fig. 3. 900×600 pixels.
- **CSK1:** HH Polarized COSMO-SkyMed, StripMap (2.5 m ground resolution), geocoded, single-look image acquired over Italy, city of Lombriasco; (©ASI, 2008); see Fig. 4. 950×700 pixels.
- **CSK2:** HH Polarized COSMO-SkyMed, StripMap (2.5 m ground resolution), geocoded, single-look image acquired over Italy, city of Cavallermaggiore; (©ASI, 2008); see Fig. 2. 650×950 pixels.

Three main classes (urban, dry soil and wet soil/water) were present in the considered images, playing important roles in the framework of flood management. We used manually annotated ground truths for training and test sets with spatially disjoint training and test fields. For all images, a small proportion of pixels (around 6%) was selected as learning samples. Given this small percentage, the β MRF parameter had to be obtained by trial and error for β values chosen in $[0.6; 2.2]$, leading to $\beta \approx 1.3$. The results are assessed both qualitatively (See Fig. 2) and quantitatively (accuracy results, Tab. 3). The CoDSEM algorithm presented in this paper was compared to the DSEM-MRF (no textural feature) and to a K -nearest neighbor (K -NN) integrated in a hidden MRF scheme.

Qualitatively, the GLCM and the semivariogram approaches lead to similar results (Tab. 3). But as we can notice in Fig.1, the semivariogram tends to generate circular artifacts in the urban area, whereas GLCM highlights more rectangular shapes, being more adapted to such areas.

Table 3 also highlights the fact that the introduction of a textural feature combined with CoDSEM increases the overall classification accuracy as compared to DSEM applied to a single SAR image, or to the K -NN-MRF algorithm applied to the combination texture/SAR amplitude. CoDSEM and K -NN-MRF may have similar behaviours (**TSX1**), thus emphasising the role of textural features. Focusing on the accuracy obtained for each class, the use of textural features allows urban and soil areas to be better classified. However, in this case, the

accuracy of water class is inferior to the one obtained by classification without integration of textural features. This can be explained by the fact that water and soil classes have quite close representations (greylevels) in the GLCM textural features (Fig. 1).

Figure 2 shows the classification maps obtained for **CSK2**, and the effects of the integration of the textural feature: urban areas are better delimited, which is an expected result by using a textural feature.

Table 3. Accuracy for each of the 3 classes and overall results for the test areas of the two COSMO-SkyMed and of the TerraSAR-X images, by using CoDSEM (image + textural feature), DSEM (only SAR image) and K -NN-MRF algorithms.

	TSX1				CSK1			
	water	urban	land	overall	water	urban	land	overall
CoDSEM (GLCM)	91.28%	98.82%	93.53%	94.54%	95.28%	98.67%	98.50%	97.48%
DSEM	92.95%	98.32%	81.33%	90.87%	97.74%	98.90%	81.80%	92.82%
K -NN-MRF	90.56%	98.49%	94.99%	94.68%	93.86%	85.54%	99.91%	93.10%

	CSK2			
	water	urban	land	overall
CoDSEM (GLCM)	98.62%	98.42%	100%	99.01%
CoDSEM (Semivar.)	98.37%	98.91%	100%	99.09%
DSEM	99.14%	98.88%	84.65%	94.22%
K -NN-MRF	96.72%	96.09%	99.92%	97.58%

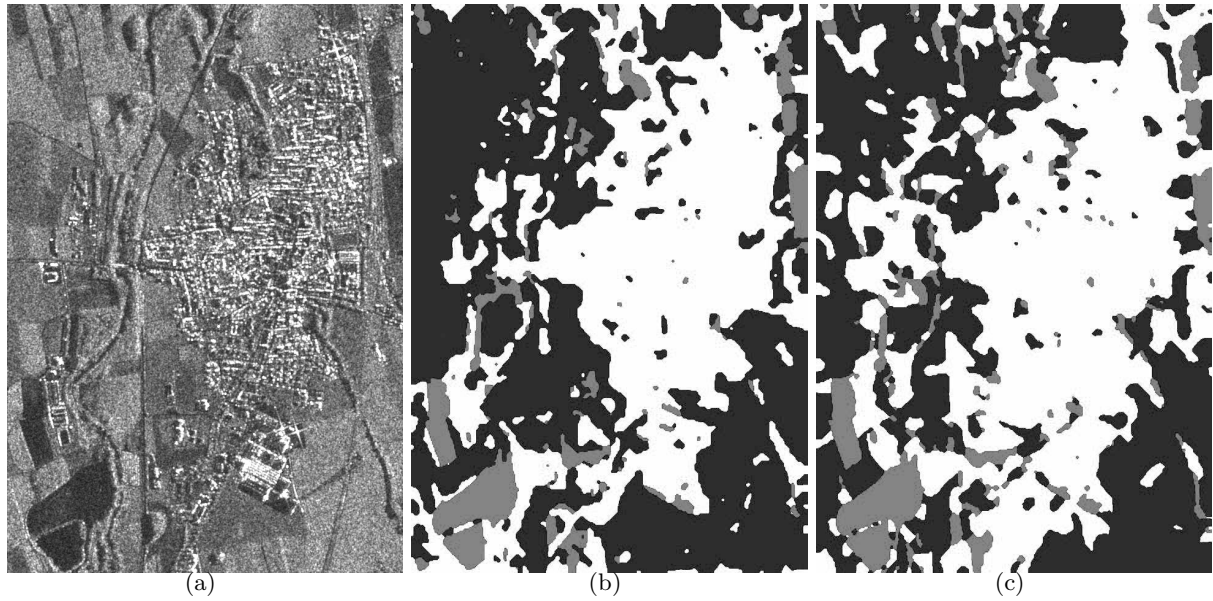


Figure 2. **CSK2**. (a): Original SAR image (COSMO-SkyMed, ©ASI); (b): Classification map obtained by Co-DSEM (with textural features obtained by GLCM); (c): Classification map obtained by DSEM (without textural features). In dark grey: land class; In middle grey: water class; In white: urban areas. The urban classification improvement is clearly observable comparing (b) and (c).

Figure 4 shows the CoDSEM classification map of **CSK1** compared to an hand-made ground truth. Globally, most of the test pixels are well classified. Erroneously classified pixels correspond to small areas, and can be explained e.g. by shadows in the city (mistaken as water due to the low greylevel value) or isolated houses in the middle of land patches and not taken into account in the manually-built ground truth.

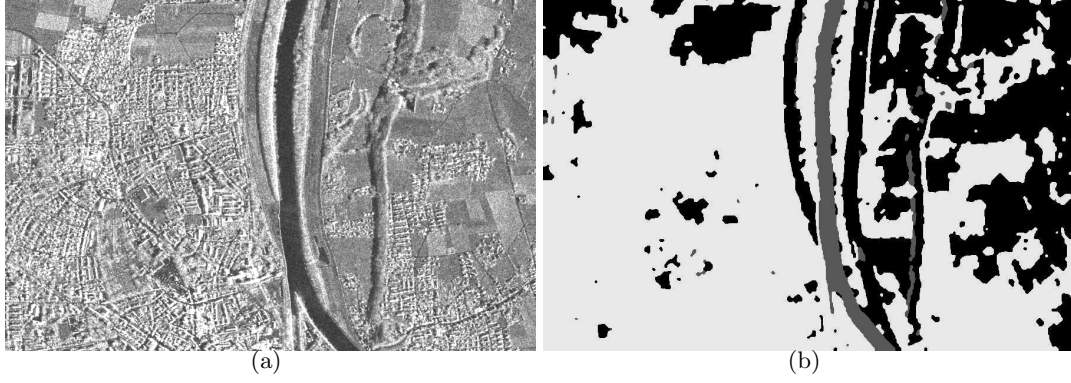


Figure 3. **TSX1**. (a): Original SAR image (TerraSAR-X,©Infoterra); (b): Classification map obtained by Co-DSEM (with textural features obtained by GLCM); In dark grey: land class; In middle grey: water class; In light grey: urban areas. Most of the pixels seems to be correctly classified. Misclassified pixels are related to the trees (e.g. shadows considered as water).

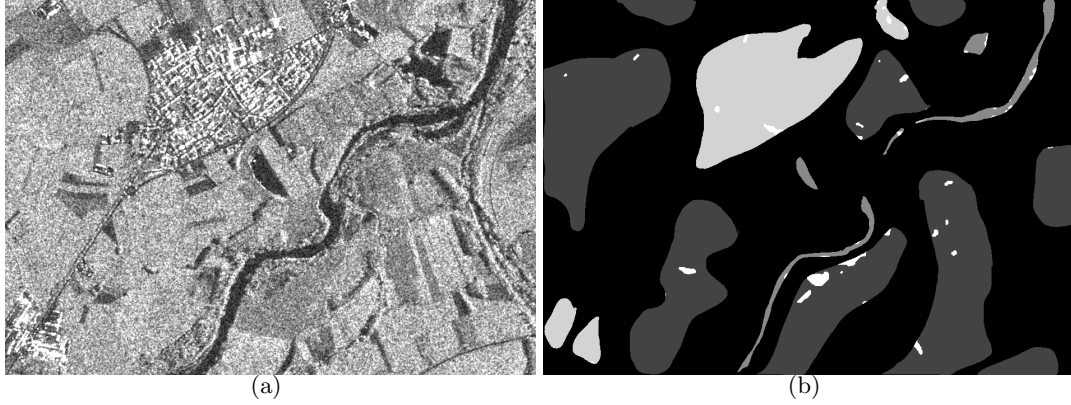


Figure 4. **CSK1**. (a): Original SAR image (COSMO-SkyMed,©ASI); (b): Classification map obtained by Co-DSEM (with textural features obtained by GLCM). In white: misclassified pixels; In black: pixels not considered in the test map; In different greylevels: match between the test ground truth and the classification label estimation (dark grey: land — middle grey: water — light grey: urban).

6. CONCLUSIONS

Urban areas are known to be socially and economically critical with respect to natural disasters. For this reason, it is specifically important to exploit VHR SAR data to obtain land-cover and land-use classification results in these areas. Already tested in wet soil classification for epidemiologic applications, the efficient CoDSEM algorithm is adapted here to the supervised classification of urban areas. First, it models the SAR image amplitude probability function via a dictionary-based stochastic expectation maximization algorithm. As a second step, it combines Bayesian classification by the Markov random field contextual approach so as to ensure the robustness of the algorithm with respect to speckle. To improve the classification results when dealing with urban areas in very high resolution, we include a textural feature as a second channel. This textural feature is extracted from the original SAR image by calculation of a GLCM texture. To take into account both SAR original image and the extracted feature, we model a joint PDF via copulas.

The accuracy of the proposed algorithm was validated in the application to urban/land/water separation on several single-pol high resolution COSMO-SkyMed images. The experiments pointed out high accuracy on all test images, but also some smoothing effects at the spatial borders between the different classes. Moreover, artifacts due to the use of a simple feature extraction are relevant. Thus, an interesting development could be

the integration of more sophisticated texture-extraction techniques. As a future work, geometrical information, relevant in urban areas, will be considered via hierarchical/multiscale MRFs.

ACKNOWLEDGMENTS

This work was carried out with the partial financial support of Institut National de Recherche en Informatique et Automatique (INRIA), France, and the Direction Générale de l'Armement (DGA), France. This support is gratefully acknowledged. The authors would like to thank the Italian Space Agency (ASI) for providing the COSMO-SkyMed images. The TerraSAR-X image of Rosenheim (©Infoterra) was taken from <http://www.infoterra.de/>, where it is available for free testing.

REFERENCES

- [1] Oliver, C. and Quegan, S., [*Understanding Synthetic Aperture Radar images*], Artech House, Norwood (1998).
- [2] Boni, G., Castelli, F., Ferraris, L., Pierdicca, N., Serpico, S., and Siccardi, F., "High resolution COSMO-SkyMed SAR data analysis for civil protection from flooding events," *IEEE Trans. Geosci. Remote Sens.*, 6–9 (2007).
- [3] Causa, F., Moser, G., and Serpico, S. B., "A spatio-spectral data fusion method for the detection of urban areas in optical high resolution images by a novel MRF model," *Proc. of IGARSS*, 2502–2505 (2006).
- [4] Lee, J.-S., Grunes, M. R., and Kwok, R., "Classification of multi-look polarimetric SAR imagery based on complex Wishart distribution," *Int. J. Remote Sensing* **15**(11), 2299–2311 (1994).
- [5] Benboudjema, D. and Pieczynski, W., "Unsupervised image segmentation using triplet Markov fields," *Computer Vision and Image Understanding* **99**(3), 476–498 (2005).
- [6] Jacob, A. M., Hemmerly, E. M., and Fernandes, D., "SAR image classification using a neural classifier based on Fisher criterion," in [*Medical Imaging: Image Processing*], *Proc. of the VII Brazilian Symposium on Neural Networks (SBRN)*, 168–172 (2002).
- [7] Yang, W., Dai, D., Triggs, B., and Xia, G.-S., "Semantic labeling of SAR images with hierarchical Markov aspect models," *HAL Research Report*, hal-00433600 (2009).
- [8] Chamundeeswari, V. V., Singh, D., and Singh, K., "Unsupervised land cover classification of SAR images by contour tracing," *IEEE Trans. Geosci. Remote Sens.*, 547–550 (2007).
- [9] Moser, G., Krylov, V., Serpico, S. B., and Zerubia, J., "High resolution SAR image classification by Markov random fields and finite mixtures," *Proc. of SPIE* **7533**, 753308 (2010).
- [10] Moser, G., Serpico, S. B., and Zerubia, J., "Dictionary-based Stochastic Expectation Maximization for SAR amplitude probability density function estimation," *IEEE Trans. Geosci. Remote Sens.* **44**(1), 188–199 (2006).
- [11] Krylov, V., Moser, G., Serpico, S. B., and Zerubia, J., "Dictionary-based probability density function estimation for high-resolution SAR data," *Proc. of SPIE* **7246**, 72460S (2009).
- [12] Nelsen, R. B., [*An introduction to copulas*], Springer, New York (2nd ed.) (2007).
- [13] Dubes, R. C. and Jain, A. K., "Random field models in image analysis," *Journal of Applied Statistics* **16**(2), 131–164 (1989).
- [14] Kato, Z., Zerubia, J., and Berthod, M., "Unsupervised parallel image classification using Markovian models," *Pattern recognition* **32**(4), 591–604 (1999).
- [15] Figueiredo, M. A. T. and Jain, A. K., "Unsupervised learning of finite mixture models," *IEEE Trans. Patt. Anal. Mach. Intell.* **24**(3), 381–396 (2000).
- [16] Besag, J., "Spatial interaction and the statistical analysis of lattice systems," *Journal of the Royal Statistical Society* **36**(2), 192–236 (1974).
- [17] Besag, J., "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society* **48**, 259–302 (1986).
- [18] Geman, S. and Geman, D., "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Patt. Anal. Mach. Intell.* **6**(6), 721–741 (1984).

- [19] Yu, Y. and Cheng, Q., “MRF parameter estimation by an accelerated method,” *Pattern Recognit. Lett.* **24**(9-10), 1251–1259 (2003).
- [20] Celeux, G., Cheveau, D., and Diebolt, J., “On stochastic versions of the EM algorithm,” *INRIA Research Report* **2514** (1995).
- [21] Tison, C., Nicolas, J.-M., Tupin, F., and Maitre, H., “A new statistical model for Markovian classification of urban areas in high-resolution SAR images,” *IEEE Trans. Geosci. Remote Sens.* **42**(10), 2046–2057 (2004).
- [22] Nicolas, J.-M. and Maruani, A., “Lower-order statistics: a new approach for probability density functions defined on R^+ ,” *Proc. of EUSIPCO*, 805–808 (2000).
- [23] Sneddon, I., [*The use of integral transforms*], McGraw-Hill, New York (1972).
- [24] Haralick, R. M., Shanmugam, K., and Dinstein, I., “Textural features for image classification,” *IEEE Trans. on Systems, Mans and Cybern.* **3**(6), 610–621 (1973).
- [25] Chen, Q. and Gong, P., “Automatic variogram parameter extraction for textural classification of the panchromatic IKONOS imagery,” *IEEE Trans. Geosci. Remote Sens.* **42**(5), 1106–1115 (2004).
- [26] Huard, D., Evin, G., and Fabre, A.-C., “Bayesian copula selection,” *Computational statistics and data analysis* **51**(2), 809–822 (2006).
- [27] Lehmann, E. and Romano, J., [*Testing statistical hypotheses*], Springer, New York (3rd ed.) (2005).